Interior-Point Methods

Wing-Kin (Ken) Ma

Department of Electronic Engineering,
The Chinese University Hong Kong, Hong Kong

Lesson 11, ELEG5481

Before we go...

- Interior-point methods play an indispensable role in convex optimization.
- Modern LP/SOCP/SDP solvers, such as SeDuMi, SDPT3, and DSDP, are interior-point methods.
- What we will do:

  Provide intuitive insights into the ideas that led to this beautiful technique.
- What we will not go through:

  Convergence/complexity analysis, coverage of all interior-point algorithms (there are numerous!), complete derivations, implementation details, . . .

Trapping Points within the Feasible Set

- Convex problem in standard form

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t.}\ & Ax = b,\ f_i(x) \le 0,\ i = 1, \ldots, m \end{aligned} \qquad (\dagger)$$

  where $f_0, \ldots, f_m$ are assumed to be convex & twice differentiable.

- The inequality constraints can be made implicit by rewriting ($\dagger$) as

$$\begin{aligned} \min\ & f_0(x) + \sum_{i=1}^{m} I_-(f_i(x)) \\ \text{s.t.}\ & Ax = b \end{aligned}$$

  where $I_-(u) = 0$ for $u \le 0$, $I_-(u) = \infty$ otherwise.

- But I_- is not differentiable.

- **The basic idea:** approximate I_- by some differentiable function.

Logarithmetic Barrier Function

- Approximate I_- by

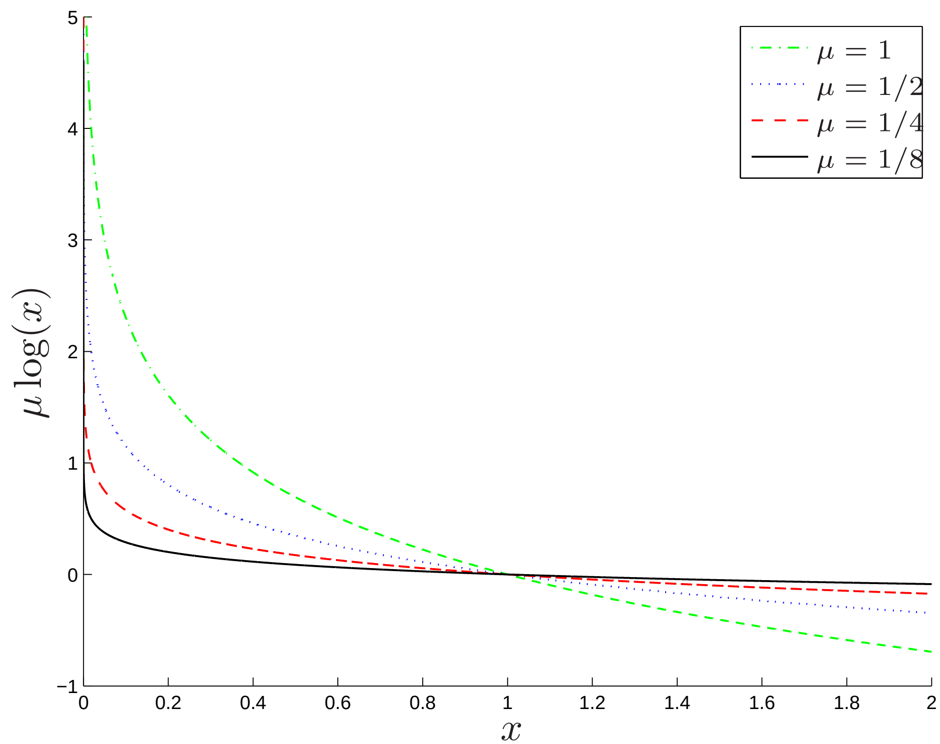
$$\hat{I}_-(u) = -\mu \log(-u), \qquad \mathbf{dom} \hat{I}_- = \{x \in \mathbf{R} \mid x < 0\}$$

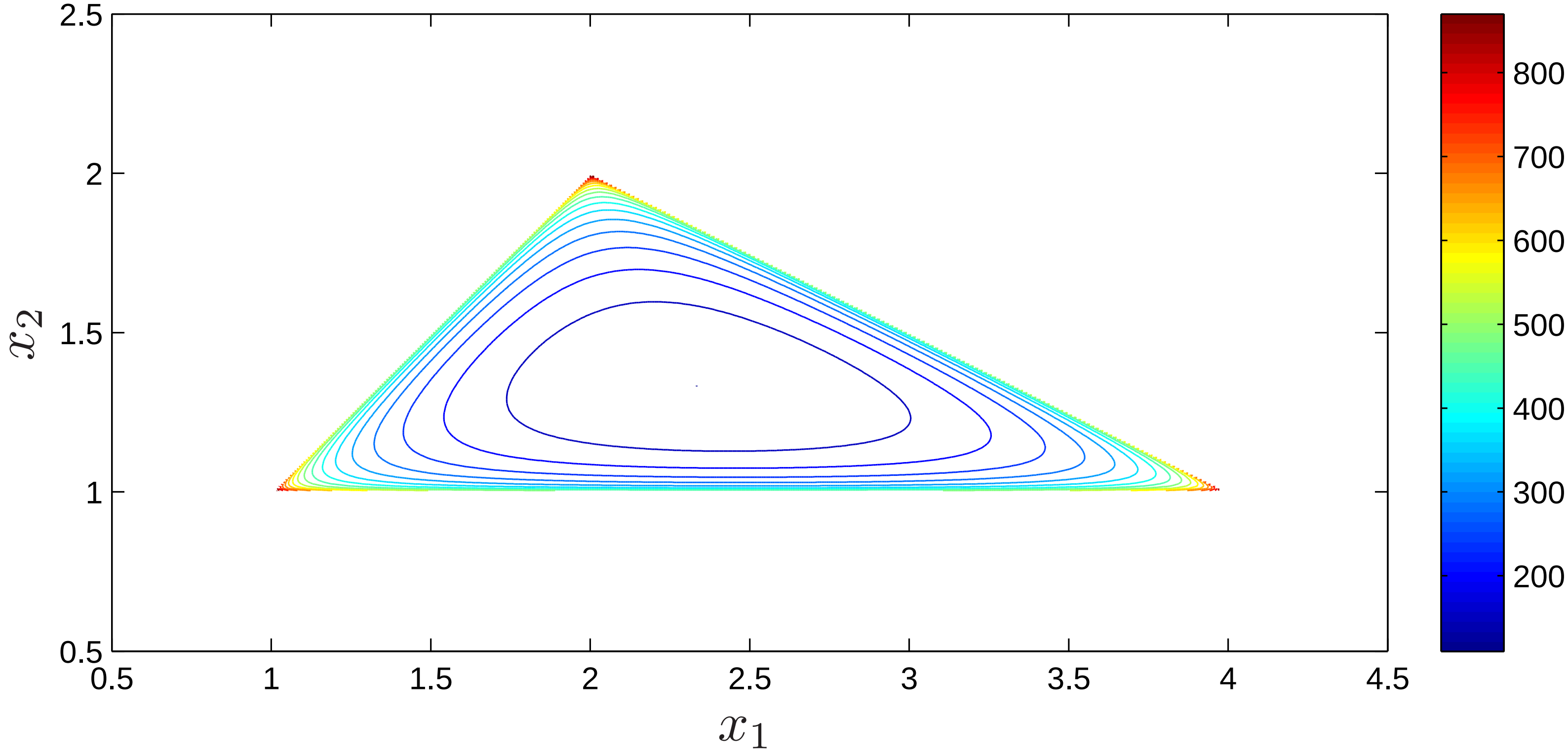
where $\mu > 0$ is a parameter that controls the accuracy of the approx.

Example: a single inequality $x \geq 0$.

The corresponding approx. is

$$-\mu \log(x)$$





Example: A set of linear inequalities $b_i - a_i^T x \leq 0$, $i = 1, \ldots, m$.

The corresponding approx. is

$$-\mu \sum_{i=1}^{m} \log(a_i^T x - b_i)$$

- An approximation of the original problem

$$\begin{aligned} \min\ & f_0(x) + \mu\phi(x) \\ \text{s.t.}\ & Ax = b \end{aligned} \tag{‡}$$

with $\phi(x) = -\sum_{i=1}^m \log(-f_i(x))$, $\textbf{dom}\phi = \{x \mid f_i(x) < 0, i = 1, \ldots, m\}$.

- ϕ is called the **logarithmetic barrier function**. Some nice properties:
  - ϕ is convex (by composition).
  - ϕ is twice differentiable:

$$\nabla\phi(x) = \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla f_i(x)$$

$$\nabla^2\phi(x) = \sum_{i=1}^m \frac{1}{f_i^2(x)} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla^2 f_i(x)$$

- This means that **the objective fn. of (‡) is convex and twice differentiable.**

Building Block: Newton's Method

- Consider unconstrained minimization of convex $f : \mathbf{R}^n \to \mathbf{R}$

$$\min \; f(x)$$

- Optimality condition: $\nabla f(x^\star) = 0$.

Newton's Method
given an initial point $x \in \mathbf{dom} f$, a tolerance $\epsilon > 0$.
repeat

1. *Compute the Newton step.*
$$\Delta x = -(\nabla^2 f(x))^{-1} \nabla f(x)$$
2. *Line search.* Choose step size α by solving
$$\min_{0 \le \alpha \le 1} f(x + \alpha \Delta x)$$
3. *Update.* $x := x + \alpha \Delta x$.
4. *Stopping criterion.* **quit** if $\frac{1}{2} \nabla f(x)^T (\nabla^2 f(x))^{-1} \nabla f(x) \le \epsilon$

- The idea: first-order approximation of the optimality condition

$$\nabla f(x + \Delta x) \approx \nabla f(x) + \nabla^2 f(x) \Delta x = 0$$

- Newton's method can be extended to handle

$$\begin{aligned} &\min \; f(x) \\ &\text{ s.t. } \; Ax = b \end{aligned}$$

  (Intuitively this is reasonable because the equality constrained min. problem is equivalent to $\min_z f(A^\dagger b + Fz)$ for some F such that $\mathcal{R}(F) = \mathcal{N}(A)$).

- Newton's method has fast convergence in general.

Central Path

- The log barrier approximation

$$\begin{aligned} \min\ & f_0(x) + \mu\phi(x) \\ \text{s.t.}\ & Ax = b \end{aligned} \qquad (*)$$

  - is an accurate approx. of the original problem for $\mu \to 0$ (with $\mu > 0$).
  - becomes difficult to minimize as $\mu \to 0$ (true at least for Newton's method)

- **The basic idea:** start with a large μ, and iteratively reduce μ until a desired solution accuracy is reached.

- Define $x^\star(\mu)$ to be the solution of $(*)$.

- **Central path**

$$\{x \mid x = x^\star(\mu), \mu > 0\}$$

  is the collection of optimal points for various μ. A property:

$$f_0(x^\star(\mu)) - p^\star \leq m\mu$$

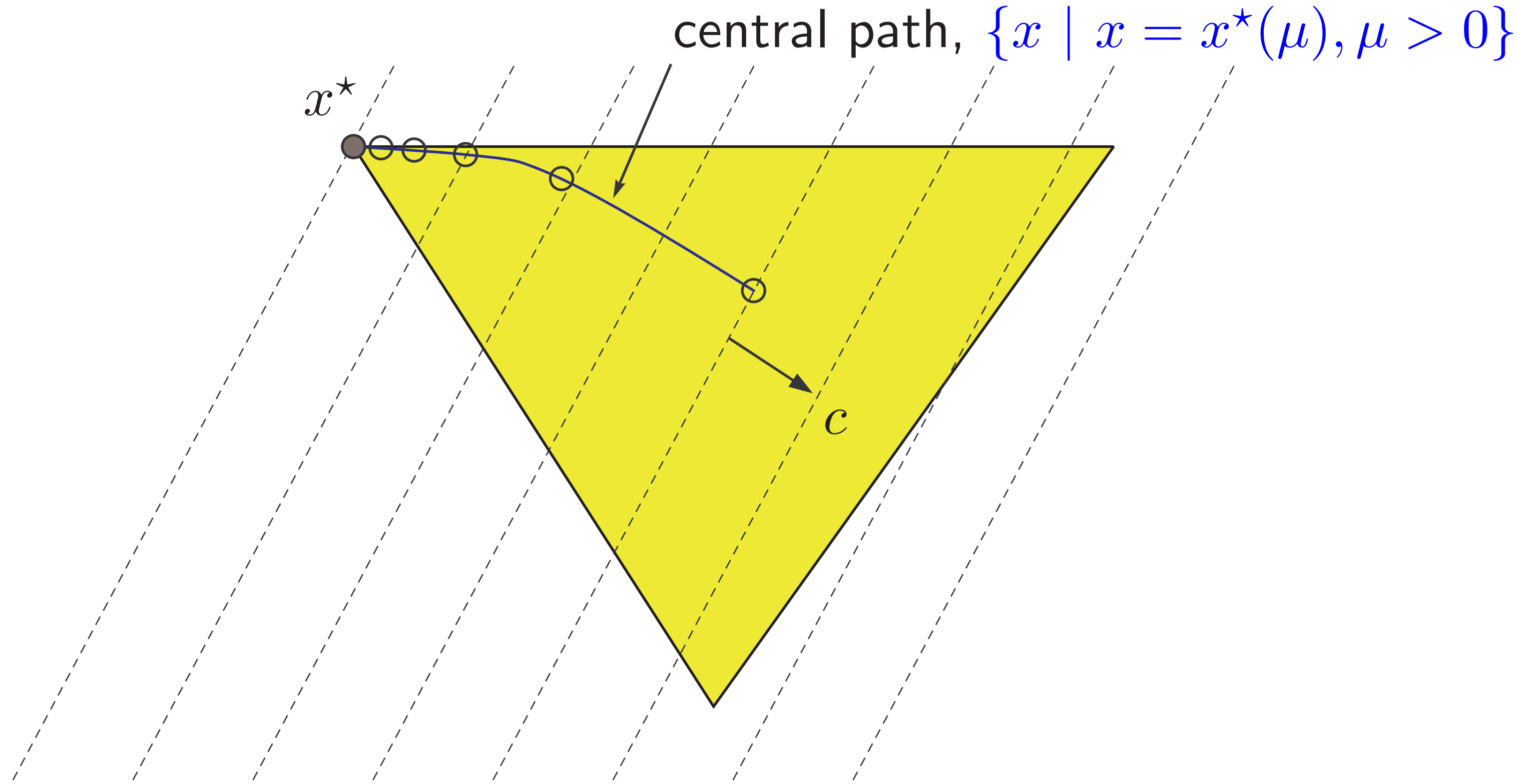


Illustration of the central path of LP. An interior-point method would follow the central path to iteratively approach the optimal solution.

A Path Following Method (or Barrier Method)

Path Following Method
given an initial strictly feasible point x, μ, $\epsilon > 0$, & $\beta < 1$.
repeat

1. *Centering step.* Starting at x, use Newton's method to solve
$$x^\star(\mu) = \min_x \mu f_0(x) + \phi(x) \text{ s.t. } Ax = b,$$
2. *Update.* $x := x^\star(\mu)$.
3. *Stopping criterion.* **quit** if $m\mu < \epsilon$.
4. *Target shifting.* $\mu := \beta\mu$.

- **Short-step path following:** choose β close to 1.
  - small number of Newton steps per outer iteration
  - large number of outer iterations
- **Long-step path following:** choose a small β.
  - increased number of Newton steps per outer iteration
  - smaller number of outer iterations

Log Barrier for Conic Optimization

- Conic problem in standard form

$$\min\ c^T x$$
$$\text{s.t. } Ax = b,\ x \succeq_K 0$$

- Log barriers for
  - **LP**: $K = \mathbf{R}^n_+$

$$\phi(x) = -\textstyle\sum_{i=1}^n \log(x_i)$$

  - **SOCP**: $K = \{(x,t) \mid \|x\|_2 \le t\}$

$$\phi(x) = \log\left(t - \|x\|_2^2/t\right)$$

  - **SDP**: $K = \{X \in \mathbf{S}^n \mid X \succeq 0\}$

$$\phi(X) = -\log\det(X)$$

- These barriers are **self-concordant** (def. skipped), from which appealing convergence results can be proven.

Primal-Dual Path Following Method for SDP

- Primal-dual path following has good control over solution accuracy.

- **The idea:** Approximate the KKT conditions of the barrier problems.

- Primal-dual path following can be applied to general convex problems, but here we use SDP as an example.

- Features of primal-dual path following

  - simultaneously produce primal and dual points at each iteration
  - the central path parameter μ is adaptively updated according to the current primal and dual points.
  - no inner-outer iterations as in primal path following

- SDP in standard form

$$\begin{aligned} \min_{X} \ & \mathbf{tr}(CX) \\ \text{s.t. } & X \succeq 0, \ \mathbf{tr}(A_i X) = b_i, \ i = 1, \ldots, m \end{aligned} \qquad \text{(SDP)}$$

where $A_i \in \mathbf{S}^n$, $b_i \in \mathbf{R}$.

- The dual problem

$$\begin{aligned} \max_{y,Z} \ & b^T y \\ \text{s.t. } & Z \succeq 0, \ C - Z - \textstyle\sum_{i=1}^m y_i A_i = 0 \end{aligned} \qquad \text{(DSDP)}$$

- The KKT conditions: for $(X, y, Z) = (X^\star, y^\star, Z^\star)$,

$$\begin{aligned} X \succeq 0, Z \succeq 0 \\ C - Z - \textstyle\sum_{i=1}^m y_i A_i = 0 \\ b_i - \mathbf{tr}(A_i X) = 0, \ i = 1, \ldots, m \\ ZX = 0 \qquad \text{(complementary slackness)} \end{aligned}$$

- Consider the barrier problem of the dual SDP (DSDP)

$$\max_{y, Z \succ 0} \; b^T y + \mu \log\det Z \qquad \text{(DBP)}$$
$$\text{s.t. } C - Z - \textstyle\sum_{i=1}^m y_i A_i = 0$$

- The KKT conditions for (DBP): for $(X, y, Z) = (X^\star(\mu), y^\star(\mu), Z^\star(\mu))$,

$$\mu Z^{-1} - X = 0$$
$$C - Z - \textstyle\sum_{i=1}^m y_i A_i = 0$$
$$b_i - \mathbf{tr}(A_i X) = 0, \; i = 1, \ldots, m$$

- A property for the duality gap:

$$\mathbf{tr}(C X^\star(\mu)) - b^T y^\star(\mu) = \mathbf{tr}(Z^\star(\mu) X(\mu)) = \mu/n$$

As $\mu \to 0$, the duality gap is zero and $(X^\star(\mu), y^\star(\mu), Z^\star(\mu)) \to (X^\star, y^\star, Z^\star)$.

Primal-Dual Path Following Method for SDP

given an initial strictly feasible point $(X, y, Z) = (X^{(0)}, y^{(0)}, Z^{(0)})$, $\epsilon > 0$.

repeat

1. *Compute the current μ.* $\mu := \mathbf{tr}(ZX)/n$.
2. *Target shifting.* $\mu := \mu/2$.
3. *Compute search directions for new μ.* Compute $(\Delta X, \Delta y, \Delta Z)$ by solving the KKT eqns.

$$\mu(Z + \Delta Z)^{-1} - (X + \Delta X) = 0$$
$$C - (Z + \Delta Z) - \textstyle\sum_{i=1}^{m} (y_i + \Delta y_i) A_i = 0$$
$$b_i - \mathbf{tr}(A_i(X + \Delta X)) = 0, \; i = 1, \ldots, m$$

in an approximate manner.

4. *Line search.* Compute primal & dual step-sizes α_p, α_d such that
$$X + \alpha_p \Delta X \succ 0 \qquad\qquad Z + \alpha_d \Delta Z \succ 0$$
5. *Update.* $X := X + \alpha_p \Delta X$, $y := y + \alpha_d \Delta y$, $Z := Z + \alpha_d \Delta Z$.
6. *Stopping criterion.* **quit** if $\mathbf{tr}(ZX) < \epsilon$.

Approximation of the KKT equations (Step 3):

For a primal-dual feasible iterate (X, y, Z), the eqns in Step 3 are expressed as

$$\mu(Z+\Delta Z)^{-1} = (X+\Delta X) \quad (1)$$

$$-\Delta Z - \textstyle\sum_{i=1}^{m} \Delta y_i A_i = 0 \quad (2)$$

$$\mathbf{tr}(A_i \Delta X) = 0, \;\; i = 1, \ldots, m \quad (3)$$

(2)-(3) are linear & are easy to handle. The difficulty mainly lies in (1).

Apply a 1st-order approximation to (1) (note: from this point there are many possible variations); e.g., in **[Helmberg *et. al*'96]**,

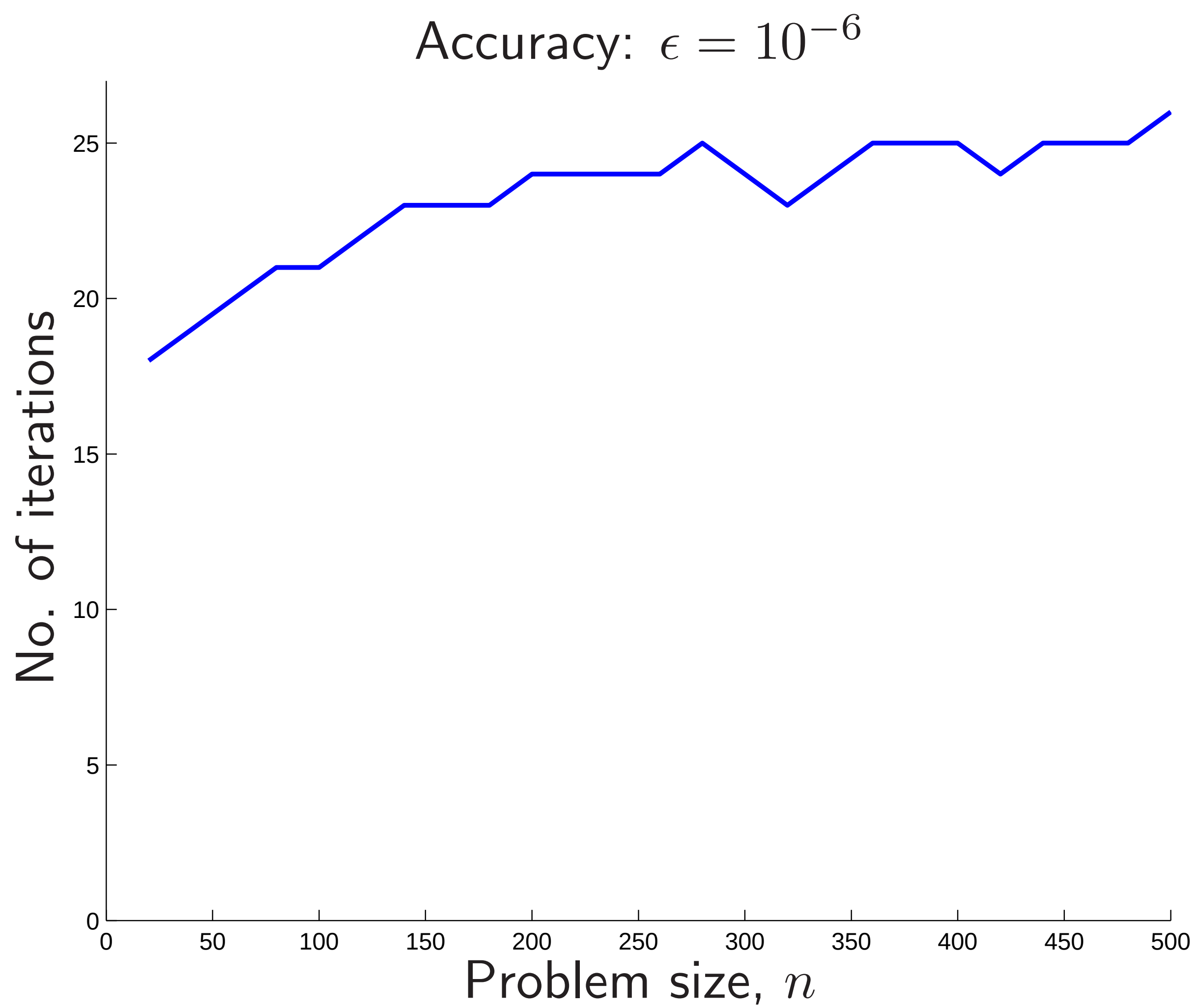
$$\begin{aligned} \mu I = (Z+\Delta Z)(X+\Delta X) &= ZX + Z\Delta X + X\Delta Z + \Delta Z \Delta X \\ &\approx ZX + Z\Delta X + X\Delta Z \quad (4) \end{aligned}$$

Equations (2)-(4) (modified linearized KKT eqns.) are linear in $(\Delta X, \Delta y, \Delta Z)$, and can be solved in closed form.

- Convergence analysis for primal-dual path following methods suggests the following:

> For a given tolerance $\epsilon > 0$ (such that $\mathbf{tr}(CX) - p^\star < \epsilon$), terminates in
>
> $$\mathcal{O}\left(\sqrt{n}\log\frac{\mathbf{tr}(X^{(0)}Z^{(0)})}{\epsilon}\right) \text{ iterations}$$
>
> in the worst case.

- Practical experience (e.g., with the MIMO detection application described later) suggests that the no. of iterations grows much more slower than $\sqrt{n}$.



Nos. of iterations of the primal-dual SDP path following method in practice. The problem is

$$\min \ \mathbf{tr}(CX) \ \text{s.t.} \ X_{ii} = 1, i = 1, \ldots, n$$

As seen, the nos. of iterations appear to be flat for large n.

Other (Even More Advanced) Interior-Point Methods

- Self-dual embedding (used in SeDuMi)
- Dual scaling (used in DSDP)
- Infeasible primal-dual path following (used in SDPT3)
- ...

References

[Boyd-Vanderberghe'04] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, 2004.

[Helmberg-Rendl-Vanderbei-Wolkowicz'96] C. Helmberg, F. Rendl, R.J. Vanderbei, and H. Wolkowicz, "An interior-point method for SDP", *SIAM J. Optim.*, 1996.